

University of Wollongong
Research Online

Faculty of Engineering and Information
Sciences - Papers: Part A

Faculty of Engineering and Information
Sciences

1-1-2012

Scene segmentation and pedestrian classification from 3-D range and intensity images

Xue Wei

University of Wollongong, xw158@uowmail.edu.au

Son Lam Phung

University of Wollongong, phung@uow.edu.au

Abdesselam Bouzerdoun

University of Wollongong, bouzer@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/eispapers>



Part of the [Engineering Commons](#), and the [Science and Technology Studies Commons](#)

Recommended Citation

Wei, Xue; Phung, Son Lam; and Bouzerdoun, Abdesselam, "Scene segmentation and pedestrian classification from 3-D range and intensity images" (2012). *Faculty of Engineering and Information Sciences - Papers: Part A*. 249.

<https://ro.uow.edu.au/eispapers/249>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Scene segmentation and pedestrian classification from 3-D range and intensity images

Abstract

This paper proposes a new approach to classify obstacles using a time-of-flight camera, for applications in assistive navigation of the visually impaired. Combining range and intensity images enables fast and accurate object segmentation, and provides useful navigation cues such as distances to the nearby obstacles and obstacle types. In the proposed approach, a 3-D range image is first segmented using histogram thresholding and mean-shift grouping. Then Fourier and GIST descriptors are applied on each segmented object to extract shape and texture features. Finally, support vector machines are used to recognize the obstacles. This paper focuses on classifying pedestrian and non-pedestrian obstacles. Evaluated on an image data set acquired using a time-of-flight camera, the proposed approach achieves a classification rate of 99.5%.

Keywords

intensity, range, 3, classification, pedestrian, images, segmentation, scene

Disciplines

Engineering | Science and Technology Studies

Publication Details

Wei, X., Phung, S. Lam. & Bouzerdoum, A. (2012). Scene segmentation and pedestrian classification from 3-D range and intensity images. IEEE International Conference on Multimedia and Expo (pp. 103-108). Australia: IEEE.

SCENE SEGMENTATION AND PEDESTRIAN CLASSIFICATION FROM 3-D RANGE AND INTENSITY IMAGES

Xue Wei, Son Lam Phung, and Abdesselam Bouzerdoum

School of Electrical, Computer and Telecommunications Engineering
University of Wollongong, Australia

xw158@uowmail.edu.au, phung@uow.edu.au, bouzer@uow.edu.au

ABSTRACT

This paper proposes a new approach to classify obstacles using a time-of-flight camera, for applications in assistive navigation of the visually impaired. Combining range and intensity images enables fast and accurate object segmentation, and provides useful navigation cues such as distances to the nearby obstacles and obstacle types. In the proposed approach, a 3-D range image is first segmented using histogram thresholding and mean-shift grouping. Then Fourier and GIST descriptors are applied on each segmented object to extract shape and texture features. Finally, support vector machines are used to recognize the obstacles. This paper focuses on classifying pedestrian and non-pedestrian obstacles. Evaluated on an image data set acquired using a time-of-flight camera, the proposed approach achieves a classification rate of 99.5%.

Index Terms— range image, intensity image, segmentation, classification, assistive navigation

1. INTRODUCTION

Most blind people use white canes to sense the world, but the cane is restricted by its short range and basic sensing capability. To obtain more useful information, such as distance, velocity, and type of obstacles, the white cane needs to be integrated with other electronic sensors. With the development of computer vision and signal processing, several navigation systems are built to assist the blind to perceive the environment. Manduchi *et al.* [1] reviewed and evaluated computer vision and assistive technology for visually impaired population.

Existing approaches to obstacle detection rely mostly on two-dimensional color images. Their major limitation is that other positioning or distance sensors are required to determine the three-dimensional (3-D) position of the obstacle. In this paper, we propose a novel obstacle detection approach that uses a 3-D range camera. This approach not only enables fast and accurate object segmentation and classification, but also estimates the distances and speeds of moving obstacles.

The paper is organized as follows. Section 2 reviews the time-of-flight camera and algorithms for processing range images. Section 3 presents the proposed obstacle detection system, including range image segmentation, feature extraction, and pedestrian versus non-pedestrian classification. Section 4 analyzes the performance of proposed methods, and Section 5 gives the concluding remarks.

2. RELATED WORK

Time-of-flight (TOF) cameras generate depth images, where each pixel represents a distance to the camera. The camera consists of LEDs to illuminate the scene, an optical filter to gather the reflected light, a CMOS image sensor to generate the depth and intensity images, and driver electronics to control the system. Compared to the other active 3-D cameras, the TOF 3-D cameras have the advantages of a large depth-of-view and a high acquisition speed [2]. Furthermore, recent TOF 3-D cameras operate in both indoor and outdoor environments by using background lighting suppression.

Numerous approaches have been proposed for color image segmentation. For range images, there are fewer methods, and they focus mainly on finding planar surfaces or regular curved surfaces. Chandrasekaran *et al.* [3] introduced a dynamic neural network to segment range images. Their method identifies the crease edge pixels, but it focuses on segmenting only eight basic surface types. An edge detection method using Laplacian operators for irregular range images was proposed by Coleman *et al.* [4]. The improved Laplacian operators reduce noise in range images and have a higher segmentation rate than the traditional Laplacian operator. Xiao and Han [5] proposed a range image segmentation based on Bayes inference and Markov random field modelling. The authors introduced the surface function parameters to group distance pixels into planar regions.

Range images provide distance and spatial features that are different from 2-D color or intensity images. For example, for *Smart Airbag* systems, Devarakota *et al.* [6] used 3-D images to classify vehicle occupants as adults or children, leaning forward or backward. A pedestrian detection

system based on depth and intensity images was proposed by Rapus *et al.* [7]. A multi-part based people detection using 2-D range data was proposed by Mozos *et al.* [8]. In their method, individual classifiers are trained to detect different body parts, and the classifiers are combined to form a final detector.

3. PROPOSED APPROACH

The proposed approach for detecting traffic obstacles in a scene has two main tasks: i) range image segmentation and ii) obstacle classification.

3.1. Segmentation method

The proposed segmentation method consists of four steps, as shown in Fig. 1. An input image is first pre-processed to reduce measurement noise. The image is then segmented into separate regions by analyzing the depth histogram. Next, the mean-shift algorithm is applied to the 3-D points to refine segmentation. Finally, post-processing is performed to reduce over-segmentation and calculate distance to objects.

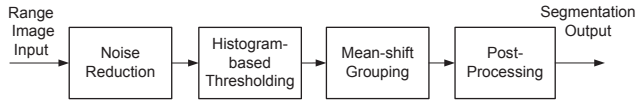


Fig. 1. Block diagram of the proposed segmentation algorithm for range images.

3.1.1. Noise pixel removal

Range cameras have higher noise and lower resolutions compared to color cameras. For outdoor environments, the output noise for range cameras is high because of the ambient light. Therefore, range cameras typically produce a confidence map to indicate the reliability of depth pixels. For example, for the SwissRanger camera, the confidence score is an integer between 0 and 7; 0 indicates the most unreliable measure and 7 indicates the most reliable measure. The confidence map is thresholded (by a value τ_1) to discard unreliable depth pixels before segmentation. Another significant artifact in range images is the salt-and-pepper noise, which can be reduced using a median filter. In the noise reduction stage, we apply both confidence map thresholding and median filtering.

In range images, object regions have the same depth values as the ground. To separate the ground from objects, we propose calculating normal vectors of surfaces. Consider the set of 3-D points (x, y, z) , where y is the vertical axis. The Delaunay triangulation is first applied on the (x, z) coordinates, which generates horizontal triangulation surface. Let

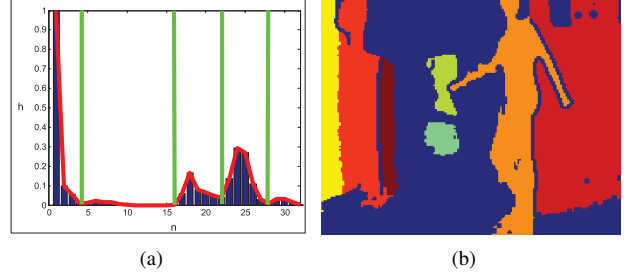


Fig. 2. Histogram-based segmentation: (a) histogram of a range image (the green lines are the detected local minima); (b) the preliminary segmentation.

\mathbf{p}_1 , \mathbf{p}_2 and \mathbf{p}_3 be the corresponding 3-D vertices of an output triangle. The normal vector of the triangulation surface is computed using the cross product:

$$(n_x, n_y, n_z) = (\mathbf{p}_3 - \mathbf{p}_2) \times (\mathbf{p}_2 - \mathbf{p}_1). \quad (1)$$

For each 3-D point, the normal vector is estimated as the average normal vectors of all triangulation surfaces at that point. A 3-D point is considered as a ground pixel if its n_y component exceeds a threshold τ_g . The processing speed to detect the ground region is slow if the entire image is processed. Since the ground region appears on the lower part of the image, the normal vectors are calculated for the bottom third region of the image.

3.1.2. Depth histogram processing

In assistive navigation, points belonging to the same object tend to have similar distances to the camera. Based on this observation, we segment image pixels into distinct distance layers using multiple thresholds. The distance thresholds are determined adaptively as the local minima of the image histogram.

Let $\mathbf{h} = \{h(1), h(2), \dots, h(N)\}$ be the image histogram with N bins. The image histogram is smoothed to reduce noise, and its first-order derivative is calculated:

$$d(n) = h(n) - h(n-1), \text{ for } n = 2, 3, \dots, N. \quad (2)$$

A local minimum is detected at point n if the first-order derivative changes its sign:

$$d(n) \times d(n-1) < 0, \text{ for } n = 3, 4, \dots, N. \quad (3)$$

After thresholding, connected component labelling is applied on each distance layer to form regions. An example of histogram-based segmentation is shown in Fig. 2. Here, the size of histogram is 32 bins.

Histogram-based segmentation has a short processing time. However, it relies only on thresholding the depth values, which leads to under-segmentation. In the next stage, 3-D data (ie. x , y , and z coordinates) are used to reduce under-segmentation.

3.1.3. Mean-shift grouping

After histogram processing, the mean-shift algorithm is applied to discover the 3-D spatial relation in under-segmented regions.

Let \mathbf{p}_i ($i = 1, \dots, n$) be a set of 3-D points that correspond to pixels in a segmented region. The mean-shift algorithm estimates the density of these 3-D points. The density centers $\hat{\mathbf{p}}$ of the 3-D points are located by kernel density estimation. This step is repeated and all under-segmented regions are partitioned into 3-D clusters. Then, the 3-D clusters are mapped to a new 2-D labelled image. The refined segmentation result is achieved.

The method to locate the centers is described as follows. Let $K(\mathbf{x})$ be a kernel function that controls how a sample \mathbf{p}_i contributes to the mean. A new estimate of the mean is calculated as

$$m(\hat{\mathbf{p}}) = \frac{\sum_{i=1}^n \mathbf{p}_i \cdot K(\hat{\mathbf{p}} - \mathbf{p}_i)}{\sum_{i=1}^n K(\hat{\mathbf{p}} - \mathbf{p}_i)}. \quad (4)$$

The difference between new mean and current mean $\{m(\hat{\mathbf{p}}) - \hat{\mathbf{p}}\}$ is the mean shift. If the mean shift is larger than a threshold $\lambda \cdot h$, the mean center is located and a new cluster is generated, where h defines the radius of kernel and λ is a scalar value. In the proposed method, $\lambda = 0.5$.

In the proposed mean-shift segmentation, the bandwidth h is adapted to the area and standard deviation of each region. The adaptive h avoids over-segmentation and reduces the computation cost. To generate the adaptive bandwidth values, the areas A_i of preliminary segmented regions are sorted. Let A be the largest area. The bandwidth h_i for the i -th segmented region is defined as

$$h_i = \frac{kA}{\delta_i A_i}, \quad (5)$$

where k is a scalar value, δ_i is the standard deviation of depth values in the region.

3.1.4. Post-processing

Two segmented regions belonging to the same object are merged by analyzing their boundaries. The common boundary of two adjacent regions \mathbf{R}_i and \mathbf{R}_j is removed if

$$\frac{W}{\min(L_i, L_j)} \leq \tau_m, \quad (6)$$

where W is the number of edge pixels on their common boundary, τ_m is a threshold, and L_i and L_j are the perimeter lengths of the two regions.

Finally, the average distance for each segmented region \mathbf{R} is calculated as

$$d = \frac{1}{N} \sum_{(i,j)} R(i, j), \quad (7)$$

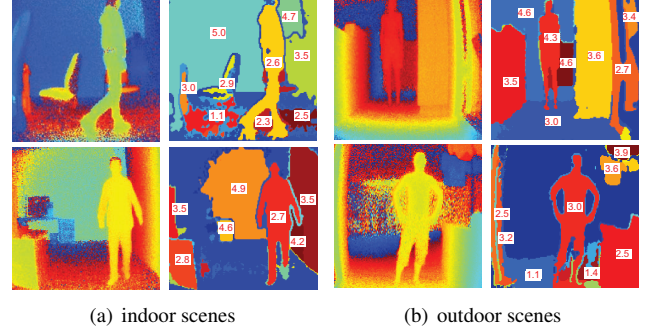


Fig. 3. The proposed segmentation method. Left: input depth images. Right: outputs. The red number is the distance in meter. See electronic color figure.

where N is number of pixels in the region. The velocity of a region is estimated as the change in distance with respect to time:

$$v = \Delta d / \Delta t. \quad (8)$$

Examples of calculated distances are shown in Fig. 3.

3.2. Obstacle classification method

For assistive navigation, the position, velocity and type of obstacles are conveyed to the blind person via speech, audio or tactile signals. To classify the segmented regions, we propose to combine the features from the range and intensity images, captured by the time-of-flight camera (see Fig. 4). For each segmented region, two types of inputs are taken from the corresponding range and intensity regions.

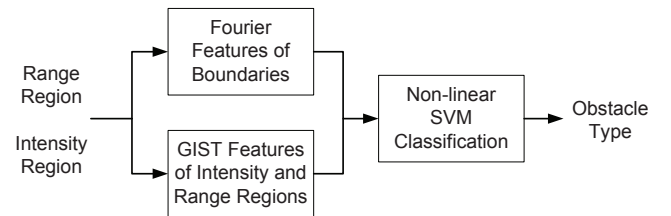


Fig. 4. Block diagram of proposed classification.

3.2.1. Fourier features from the boundary

Fourier descriptors (FD) are widely used for shape analysis. The lower frequencies represent rough shapes of objects and the higher frequencies contain finer features of the shape. To calculate Fourier coefficients, the segmented region is filtered first to remove noise pixels. Then the filtered region is thresholded into a binary region containing only the main object (shown in Fig. 5c). From the binary image, the closed boundary is traced clockwise within 8 neighbor pixels. The coordinates of the n th pixel on the boundary is represented by a

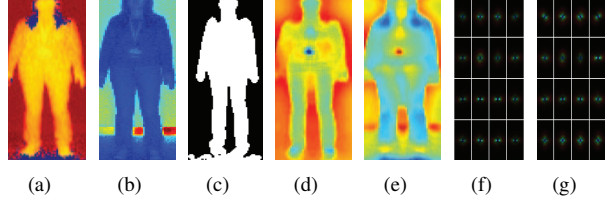


Fig. 5. Pedestrians in (a) range region, (b) intensity region, (c) boundary tracing region, (d) pre-filtered range region, (e) pre-filtered intensity region, (f) GIST magnitude of the range image on polar plot and (g) GIST magnitude of the intensity image on polar plot.

complex number. It is formed as $p_n = x_n + j y_n$, where x_n and y_n are the pixel coordinates in x-axis and y-axis.

The discrete Fourier transform is applied to the complex vectors of boundary coordinates. To obtain Fourier coefficients that are invariant with respect to scale, translation and rotation, the following steps are used.

- Fourier coefficients are calculated from segmented regions with different aspect ratios.
- For scale invariance, the Fourier coefficients F are normalized against the amplitude peak (the magnitude of the first Fourier coefficient):

$$F = \frac{F}{|F_1|}. \quad (9)$$

- For translation invariance, the first Fourier coefficient is eliminated:

$$F = \{F_2, F_3, \dots, F_n\}. \quad (10)$$

- For rotation invariance, the phases of Fourier coefficients are discarded and only the magnitudes are used:

$$F = |F|. \quad (11)$$

3.2.2. GIST features from range and intensity regions

The GIST descriptor is a holistic and low dimensional representation of blurred images [9]. It is formed from the outputs of multi-scale oriented Gabor filters. For each segmented region, we apply the GIST descriptor to the corresponding rectangular regions in the range and intensity images.

1. First, an input region is padded, whitened and normalized to reduce the blocking artifact (see Fig. 5d and e).
2. A set of multi-scale oriented Gabor filters with four scales and eight orientations is applied to the region:

$$O(x, y) = I(x, y) \otimes g(x, y). \quad (12)$$

The Gabor filters are generated from one mother wavelet, through dilation and rotation. The impulse response of a Gabor filter is the product of a harmonic function and a Gaussian function:

$$g(x, y) = \cos(2\pi \frac{x'}{\lambda} + \Phi) \exp\{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\}, \quad (13)$$

where $x' = x \cos \theta + y \sin \theta$ and $y' = -x \sin \theta + y \cos \theta$.

3. The output of each filter is partitioned into 16 blocks (see Fig. 5f and g). For each block, an average value is calculated:

$$f = \frac{1}{m \cdot n} \sum_{x=1}^m \sum_{y=1}^n |O(x, y)|. \quad (14)$$

Therefore, each filter produces 16 features.

4. The features produced by all 4×8 Gabor filters are used for object classification. The GIST features have been shown to be invariant to scale, orientation, and aspect ratio of objects [9].

The proposed feature vector consists of 512 GIST features from the range region, 512 GIST features from the intensity region, and 100 Fourier features from boundaries. We apply the SVM with the RBF kernel to discriminate pedestrian from non-pedestrian obstacles. The RBF kernel transforms input features into a high dimensional feature space:

$$K(x_i, x_j) = \exp\{-\gamma \|x_i - x_j\|^2\}, \quad (15)$$

where γ is a positive scalar.

4. EXPERIMENTS AND RESULTS

In this section, we analyze the performance of proposed methods on a data set comprising 3-D range images and intensity images.

4.1. Image data and performance measures

The data set was acquired using a TOF camera produced by MESA Imaging (model SwissRanger 4000). The frame rate is 30 frames per second and the frame size is 144×176 pixel. The depth of field of this TOF camera is 5m. The images were taken in indoor and outdoor environments, on different days, in various lighting conditions, and at different scenes. For each pixel, the camera produces five outputs: x, y, and z coordinates, amplitude, and the confidence map.

The segmentation performance is evaluated based on the Jaccard coefficient:

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad (16)$$

Table 1. Segmentation performances of the six algorithms.

Algorithms	Correct segmentation rate (%)	False segmentation rate (%)	Average Jaccard (%)	Weighted Jaccard (%)	Processing time (s)
Proposed method	73.1 ± 13.0	43.2 ± 16.1	52.7 ± 11.9	59.6 ± 13.0	0.8
Local variation [10]	64.8 ± 13.6	48.7 ± 16.8	50.4 ± 11.5	56.1 ± 12.6	3.4
Markov random field [11]	59.3 ± 17.6	48.9 ± 17.7	49.6 ± 10.6	55.9 ± 13.9	30.0
Graph-cut [12]	42.3 ± 29.0	75.7 ± 19.4	35.1 ± 9.8	40.9 ± 11.8	0.5
Otsu	41.9 ± 31.8	77.9 ± 18.7	33.6 ± 10.4	39.8 ± 12.7	0.1
K-means	39.6 ± 29.7	77.4 ± 18.8	33.3 ± 10.2	39.2 ± 11.6	0.4

where A is the ground-truth region and B is a machine segmented region. A region is considered correctly segmented if a machine segmented region exists with a high Jaccard coefficient.

Consider an image with m ground-truth regions and n machine-segmented regions. Let c be the number of correctly segmented regions by the machine. Let f be the falsely segmented regions by the machine. The correct segmentation rate P_c and false segmentation rate P_f for the image are defined as

$$P_c = c/m, \quad (17)$$

$$P_f = f/n. \quad (18)$$

The segmentation quality is measured by the average Jaccard coefficient for all ground-truth regions:

$$J_a = \frac{1}{m} \sum_{i=1}^m J_i \quad (19)$$

The weighted average Jaccard coefficient that takes into account the region size is defined as

$$J_w = \frac{\sum_{i=1}^m J_i \times A_i}{\sum_{i=1}^m A_i}, \quad (20)$$

where A_i is the area of ground-truth region i .

4.2. Analysis of image segmentation

There are two parameters in the histogram-based thresholding: the threshold of confidence map τ_1 and the histogram size N . With a lower τ_1 , unreliable pixels reduce the segmentation accuracy. With a higher τ_1 , some object regions with reliable pixels are removed. For histogram size $N = 8$, images are segmented into larger regions and under-segmentation occurs. For $N = 32$, most objects are segmented properly. For $N = 64$, images are over-segmented and post-processing cannot recover all regions correctly. In the experiments, $\tau_1 = 2$ and $N = 32$. For ground region removal, parameter τ_g is set to 0.3; for region merging, parameter τ_m is set to 0.2.

Overall, 200 range image were collected for evaluating segmentation performance. We compared the proposed algorithm for range image segmentation with several methods:

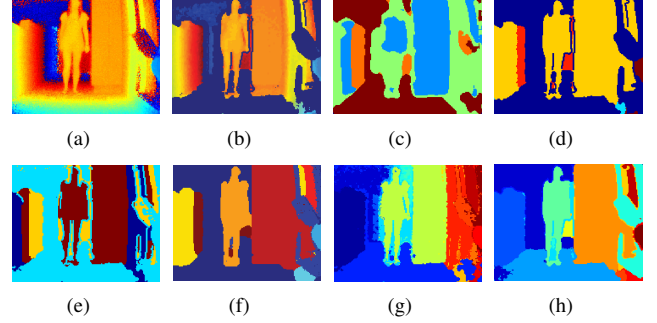


Fig. 6. Visual results of segmentation: (a) input image, (b) filtered image, (c) K-means, (d) Otsu, (e) graph-cut, (f) Markov random field, (g) local variation, (h) proposed algorithm. See electronic color figure.

Markov random field [11], local variation [10], graph cut [12], K-means and Otsu's method. The same post-processing steps are applied to all methods. The results in terms correct segmentation rate, false segmentation rate, average and weighted average Jaccard coefficients, and processing time are summarized in Table 1. Compared to the above methods, the proposed algorithm has a higher correct segmentation rate and a lower false segmentation rate. It also has higher Jaccard coefficients. The visual results of the six algorithms on a test image are shown in Fig. 6.

4.3. Analysis of pedestrian classification

In the feature extraction stage, Fourier and GIST features in range and intensity images are combined to distinguish pedestrian and non-pedestrian regions. Several state-of-the-art feature extractors are also compared using the MESA image database. In this experiment, the Fourier and GIST features were evaluated on a set of 1000 range patterns and 1000 intensity patterns. There are 23 subjects included in the dataset. The background is also varied to include both indoor and outdoor scenes. The classification rate was measured by ten-fold cross validation. For each validation fold, 9 subsets were used for training, and the remaining subset was used for testing. This was repeated ten times (for 10 choices of the test subset), and the classification rate was averaged. Visual results of the

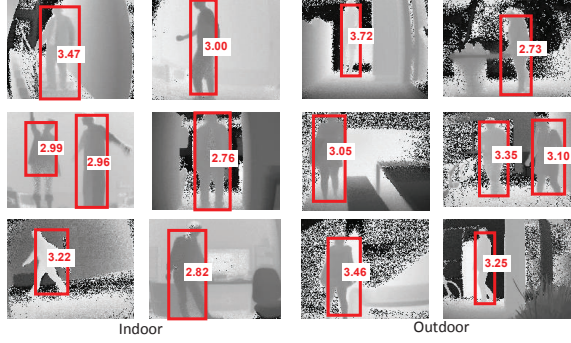


Fig. 7. Visual results of the proposed pedestrian classification.

Method	Classification Rate (%)
Fourier	90.1 \pm 2.5
HOG	91.6 \pm 3.0
SIFT	97.5 \pm 0.6
GIST	98.2 \pm 1.8
Proposed (GIST + Fourier)	99.5 \pm 0.3

Table 2. Performance of pedestrian versus non-pedestrian classification

pedestrian classification method are shown in Fig. 7. The red number is the distance in meters from the pedestrians.

For comparison, three other feature extractors (SIFT, HOG, and Fourier) were evaluated on the same data set. The results are shown in Table 2, and the main findings can be summarized as follows.

- The Fourier descriptor has a classification rate of 90.1%. The Fourier features rely on the accuracy of boundary tracing, and it only extracts shape features of objects. The noise of the boundary affects the classification rate.
- The HOG descriptor has a classification rate of 91.6%. The histogram of orientations in low-resolution range images did not perform well as in color images. All segmented regions are reshaped to a fixed aspect ratio, which increases the false classification.
- The SIFT descriptor has a classification rate of 97.5%. The SIFT is suitable for recognizing objects based on texture. However, in range images object texture is not so dominant unless there is a significant change in distance. All segmented regions are also reshaped to a fixed aspect ratio.
- The proposed method has the highest classification rate (99.5%). A possible reason is that the GIST method extracts global texture information in range and intensity images, and ignores details in the inner regions of

objects. Fourier features enhance the boundary features of objects. For low-resolution range images, shape and contour are the most dominant features.

5. CONCLUSION

In this paper, we have presented a new approach for detecting obstacles using a time-of-flight range camera. We proposed an image segmentation method that reduces over segmentation by combining histogram processing, mean-shift grouping and surface normal vectors. In this paper, we have also proposed obstacle classification algorithm by combining Fourier features and GIST features in range and intensity images. Compared with other feature extractors: Fourier descriptor, HOG and SIFT, the proposed method is found to be more effective in classifying pedestrian versus non-pedestrian using a range camera. We plan to extend the proposed system to classify other obstacles such as cars and cyclists.

6. REFERENCES

- [1] R. Manduchi and J. Coughlan, “(computer) vision without sight,” *Commun. ACM*, vol. 55, no. 1, pp. 96–104, 2012.
- [2] PolyWorks, “3-D metrology hardware review,” 2010, <http://www.innovmetric.com>.
- [3] V. Chandrasekaran, M. Palaniswami, and Terry M. Caelli, “Range image segmentation by dynamic neural network architecture,” *Pattern Recognition*, vol. 29, no. 2, pp. 315–329, 1996.
- [4] S. A. Coleman, B. W. Scotney, and S. Suganthan, “Edge detecting for range data using laplacian operators,” *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2814–2824, 2010.
- [5] X. Wang and H. Wang, “Markov random field modeled range image segmentation,” *Pattern Recognition Letters*, vol. 25, no. 3, pp. 367–375, 2004.
- [6] P. R. Devarakota, M. Castillo-Franco, R. Ginhoux, B. Mirbach, and B. Ottersten, “Occupant classification using range images,” *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1983–1993, 2007.
- [7] M. Rapus, S. Munder, G. Barattoff, and J. Denzler, “Pedestrian recognition using combined low-resolution depth and intensity images,” in *IEEE Intelligent Vehicles Symposium*, 2008, pp. 632–636.
- [8] O. M. Mozos, R. Kurazume, and T. Hasegawa, “Multi-part people detection using 2D range data,” *International Journal of Social Robotics*, vol. 2, no. 1, pp. 31–40, 2010.
- [9] A. Oliva and A. Torralba, “Modeling the shape of the scene: A holistic representation of the spatial envelope,” *International Journal of Computer Vision*, vol. 42, no. 3, pp. 145–175, 2001.
- [10] P. F. Felzenszwalb and D. P. Huttenlocher, “Efficient graph-based image segmentation,” *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.
- [11] O. Demirkaya, M. H. Asyali, and P. Sahoo, *Image Processing With MATLAB: Applications in Medicine And Biology*, CRC Press, 2008.
- [12] Y. Boykov and V. Kolmogorov, “An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.